Genomics 93 (2009) 52–61

Contents lists available at ScienceDirect

Genomics

journal homepage: www.elsevier.com/locate/ygeno

In vitro pharmacogenomic database and chemosensitivity predictive genes in gastric cancer

Jae-Joon Jung [a,b], Hei-Cheul Jeung [a,c,e], Hyun Cheol Chung [a,b,c,e], Jung Ok Lee [a,c], Tae Soo Kim [a,c], Yong Tai Kim [c], Sung Hoon Noh [a,b,d], Sun Young Rha [a,b,c,e,*]

[a] *Cancer Metastasis Research Center, Yonsei University College of Medicine, Seoul, 120-752, Korea*
[b] *Brain Korea 21 Project for Medical Science, Yonsei University College of Medicine, Seoul, 120-752, Korea*
[c] *Yonsei Cancer Center, Yonsei Cancer Research Institute, Yonsei University College of Medicine, Seoul, 120-752, Korea*
[d] *Department of Surgery, Yonsei University College of Medicine, Seoul, 120-752, Korea*
[e] *Department of Internal Medicine, Yonsei University College of Medicine, Seoul, 120-752, Korea*

ARTICLE INFO

Article history:
Received 13 May 2008
Accepted 4 August 2008
Available online 22 October 2008

Keywords:
Pharmacogenomic database
Chemotherapy
Predictive marker
Gastric cancer

ABSTRACT

Gastric cancer is one of the most common cancers worldwide, and there are clinical caveats in predicting tumor response to chemotherapy. This study describes the construction of an in vitro pharmacogenomic database, and the selection of genes associated with chemosensitivity in gastric cancer cell lines. Gene expression and chemosensitivity databases were integrated using the Pearson correlation coefficient to give the GC-matrix. The 85 genes were selected that were commonly associated with chemosensitivity of the major anticancer drugs. We then focused on the genes that were highly correlated with each specific drug. Classification of cell lines based on the set of genes associated with each drug was consistent with the division into resistant or sensitive groups according to the chemosensitivity results. The GC-matrix of the gastric cancer cell line database was used to identify different sets of chemosensitivity-related genes for specific drugs or multiple drugs.

Introduction

Chemotherapy is one of the most important modalities of cancer treatment, and many investigators have tried to improve its efficacy and reduce its toxicity. However, there is still insufficient understanding of the targets of drug activity and individual variability of patients. These limitations are in part due to a lack of effective predictive markers for drug sensitivity or resistance, and the clinical application of the few predictive markers that have been identified is not successful [1–6]. A number of genes have been reported to influence chemosensitivity. However, it is clear that the sensitivity of cancer cells to particular anticancer drugs cannot be predicted by a single factor, but is determined by many factors that influence overall sensitivity. To establish the optimal prediction system, new techniques are being applied to identify sets of chemosensitivity-related genes that could pharmacogenomically characterize the response of cancer cells to the particular anticancer drugs.

Gastric cancer is one of the most common cancers worldwide, and being treated using many chemotherapeutic agents. Although advanced gastric cancer is generally less responsive to chemotherapy, newly developed agents and combination chemotherapy have improved tumor responses. Several studies have reported that chemotherapy for gastric cancer increased not only the survival of patients but also the quality of life [7–9]. However, there are still clinical caveats in predicting tumor response in gastric cancer.

Microarray technology has been used to screen genes that are associated with chemosensitivity in cancer cell lines [2–5]. This technology has facilitated the analysis of genome-wide expression profiles that can efficiently generate information on a large scale in clinical or biological samples. More recently, spotted oligonucleotide microarrays, with a 70-base length, are used in genome screening instead of cDNA microarrays. This technology provides a high quality result, avoid clone validation, tracking and maintenance, and minimize cross-hybridization [10–16].

In this study, we constructed an *in vitro* G-matrix (gene expression database) using 22K human oligo chips, and a C-matrix (chemosensitivity database) from 13 gastric cancer cell lines treated with 16 anticancer agents. To complete the pharmacogenomic database, GC-matrixes were produced from integrating the G- and C-matrices using the Pearson correlation coefficient. The GC-matrix was initially used to select common chemosensitivity-related genes, then to identify genes that correlated strongly with the response to single anticancer agent, which would be predictive markers of chemosensitivity for each drug.

* Corresponding author. Division of Medical-Oncology, Department of Internal Medicine, Yonsei Cancer Center, Yonsei University College of Medicine, Seodaemun-Ku, Shinchon-Dong 134, Seoul, 120-752, Korea. Fax: +82 2 393 3652.
E-mail address: rha7655@yuhs.ac (S.Y. Rha).

0888-7543/$ – see front matter © 2008 Elsevier Inc. All rights reserved.
doi:10.1016/j.ygeno.2008.08.002

Results

Gene expression profiling of gastric cancer cell lines (G-matrix)

To investigate the genetic characteristics of each cell line, we performed gene expression profiling of 13 gastric cancer cell lines without drug treatment. A total of 17,015 genes that satisfy the NMP 100% among the 13 cell lines were selected for further analysis (G-matrix, Supplementary data file 1). When unsupervised hierarchical clustering was done using $\log_2$ (Red/Green) ratio, there was no significant grouping of cell lines based on anatomical origin or ethnicity (Fig. 1a). Note that the triplicates of YCC-10 clustered tightly together than any other cell lines.

Chemosensitivity profiling of gastric cancer cell lines (C-matrix)

Sixteen anticancer drugs were tested for growth inhibition in 13 gastric cancer cell lines. When unsupervised hierarchical clustering was performed with C-matrix composed of mean centered $\log_{10}IC_{50}$ (Fig. 1b, C-matrix, Supplementary data file 2), the drugs were classified into several groups according to their mechanism of action, including tubulin inhibitors (paclitaxel, docetaxel, vinblastin, and vincristine), topoisomerase I inhibitors (irinotecan and topotecan), and DNA alkylators (carboplatin and cisplatin). Red denotes high IC_{50} (resistant) and green denotes low IC_{50} (sensitive). As expected, the IC_{50}s of leucovorin and cyclophosphamide in all cell lines were higher than the Cmax of each drug. In addition, the IC_{50}s of methotrexate were also higher than its Cmax (data not shown). We observed that the clustering pattern of the 13 gastric cancer cell lines was different from that of G-matrix, as the cell lines clustered according to similarity of gene expression in the G-matrix, while clusters were arranged by drug sensitivity in C-matrix.

Correlation between gene expression and chemosensitivity (GC-matrix)

To screen the genes associated with chemosensitivity, the two databases for gene expression and chemosensitivity were integrated into one database matrix (GC-matrix, Supplementary data file 3) using the Pearson correlation coefficient. Using the GC-matrix, a hierarchical clustering was performed based on the correlation between drug sensitivity and gene expression for each drug (Fig. 1c). There were slight discrepancies in the clustering of drugs in the GC-matrix compared to the C-matrix. For example, 5-FU, one of the antimetabolites, was clustered with tubulin

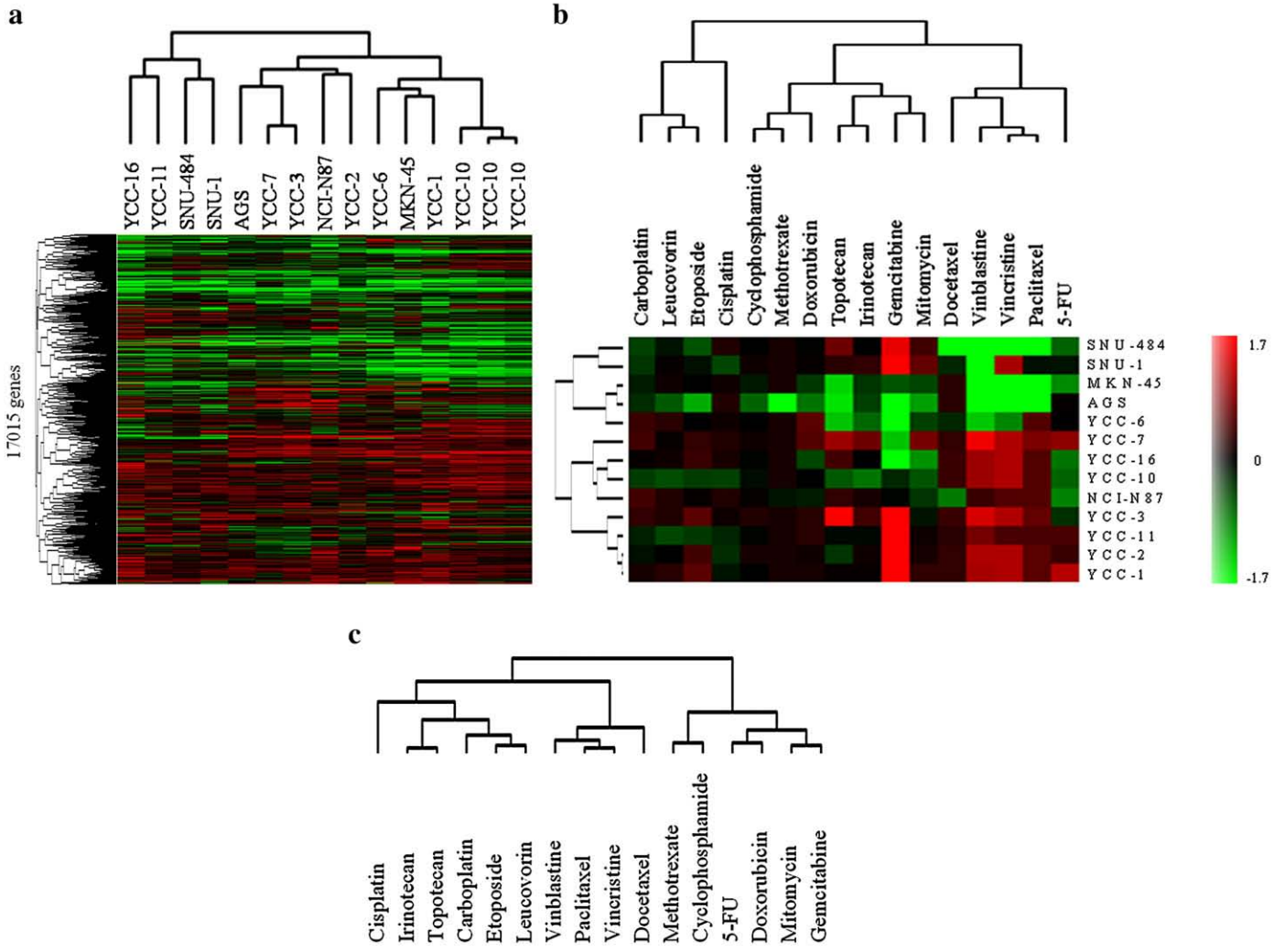

Fig. 1. G- and C-matrices. (a) G-matrix: unsupervised hierarchical clustering of 13 gastric cancer cell lines based on the $\log_2$(Red/Green) ratio of 17015 genes (NMP 100%). The triplicates of YCC-10 clustered tightly together and were well differentiated from even the most closely related cell lines. (b) C-matrix: unsupervised hierarchical clustering of 13 gastric cancer cell lines based on the $\log_{10}IC_{50}$ of 16 drugs. The scale bar on the right side of treeview indicates relative drug sensitivity; red denotes high IC_{50} (resistant) and green denotes low IC_{50} (sensitive). (c) Drug cluster of GC-matrix: unsupervised hierarchical clustering of 17015 genes and 16 drugs based on the Pearson correlation coefficient.

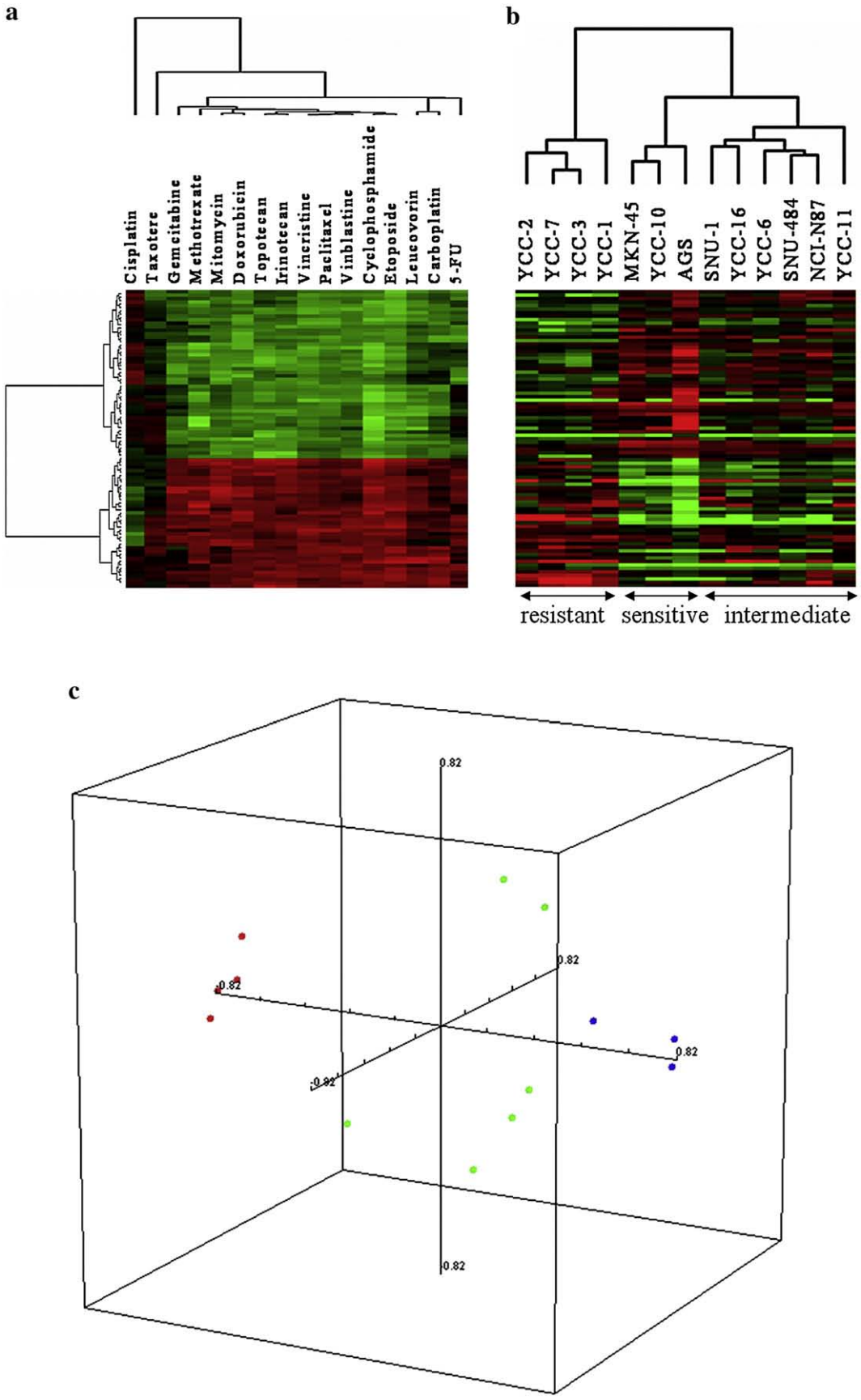

Fig. 2. Genes associated with common chemosensitivity. (a) Unsupervised hierarchical clustering of 85 common chemosensitivity-related genes and 16 drugs using their Pearson correlation coefficient. Green represents negative correlation (chemo-sensitive), and red represents positive correlation (chemo-resistant). (b) Unsupervised hierarchical clustering of 85 genes and 13 gastric cancer cell lines based on their $\log_2$(R/G) ratio. (c) Multidimensional scaling of the expression of 85 genes. Common chemo-resistant cell lines,-sensitive cell lines, and intermediate cell lines were displayed in red, blue, and green circles, respectively.

inhibitors in the C-matrix, but was correctly clustered with other antimetabolites in the GC-matrix. However, the drugs remained clustered based on their known mechanism of action, such as topoisomerase I inhibitors, tubulin inhibitors, and antimetabolites.

Genes associated with common chemosensitivity

Using the GC-matrix, we identified 85 genes which had correlation coefficients over |0.40| in 11 of 16 anticancer drugs. These 85 genes are

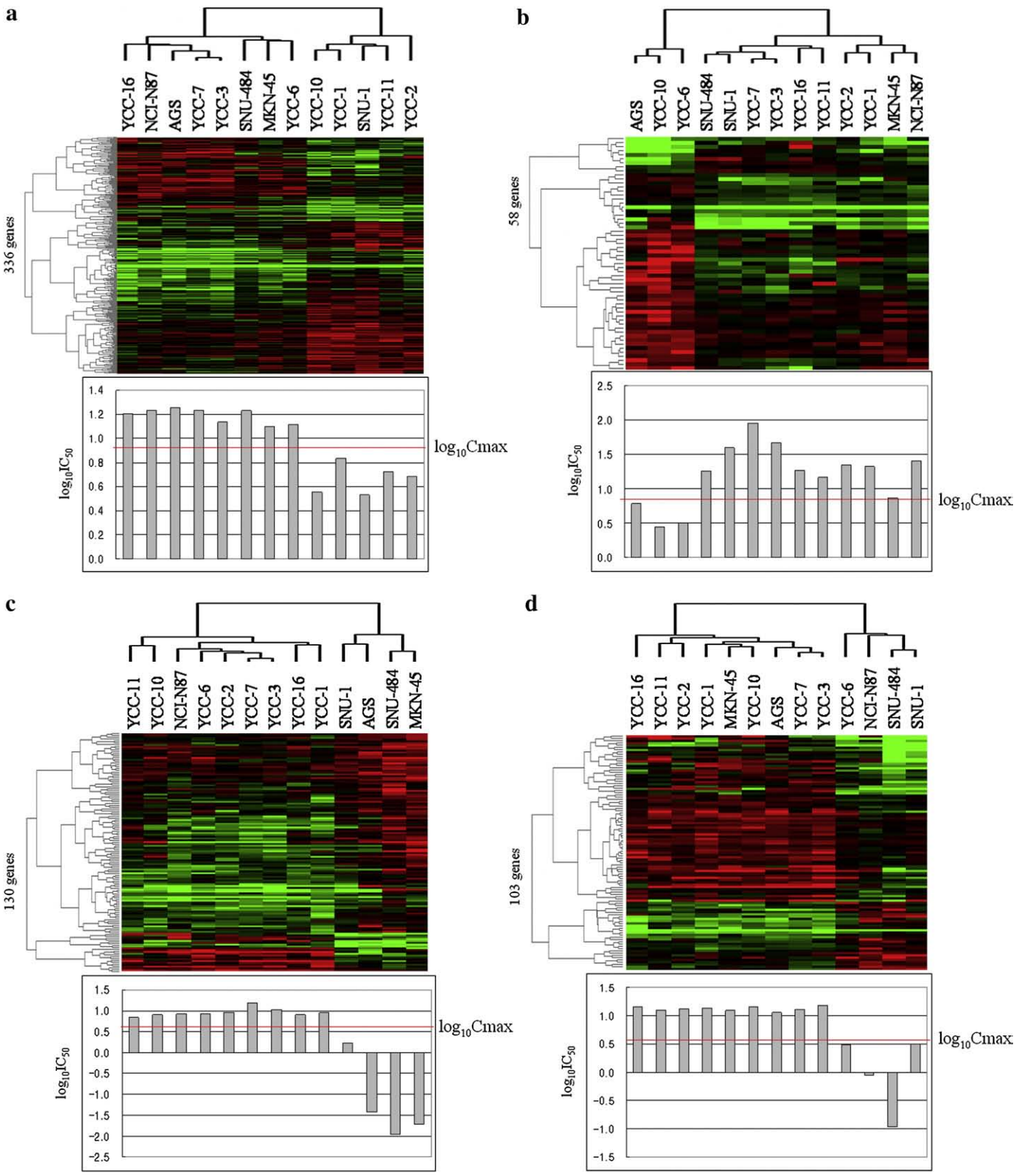

Fig. 3. Supervised hierarchical clustering with selected genes and gastric cancer cell lines and their chemosensitivity patterns. (a) Cisplatin (FDR = 0). (b) Irinotecan (FDR = 0.139). (c) Paclitaxel (FDR = 0). (d) Doxetaxel (FDR = 0.0046). (e) 5-FU (FDR = 0). (f) Doxorubicin (FDR = 0.19).

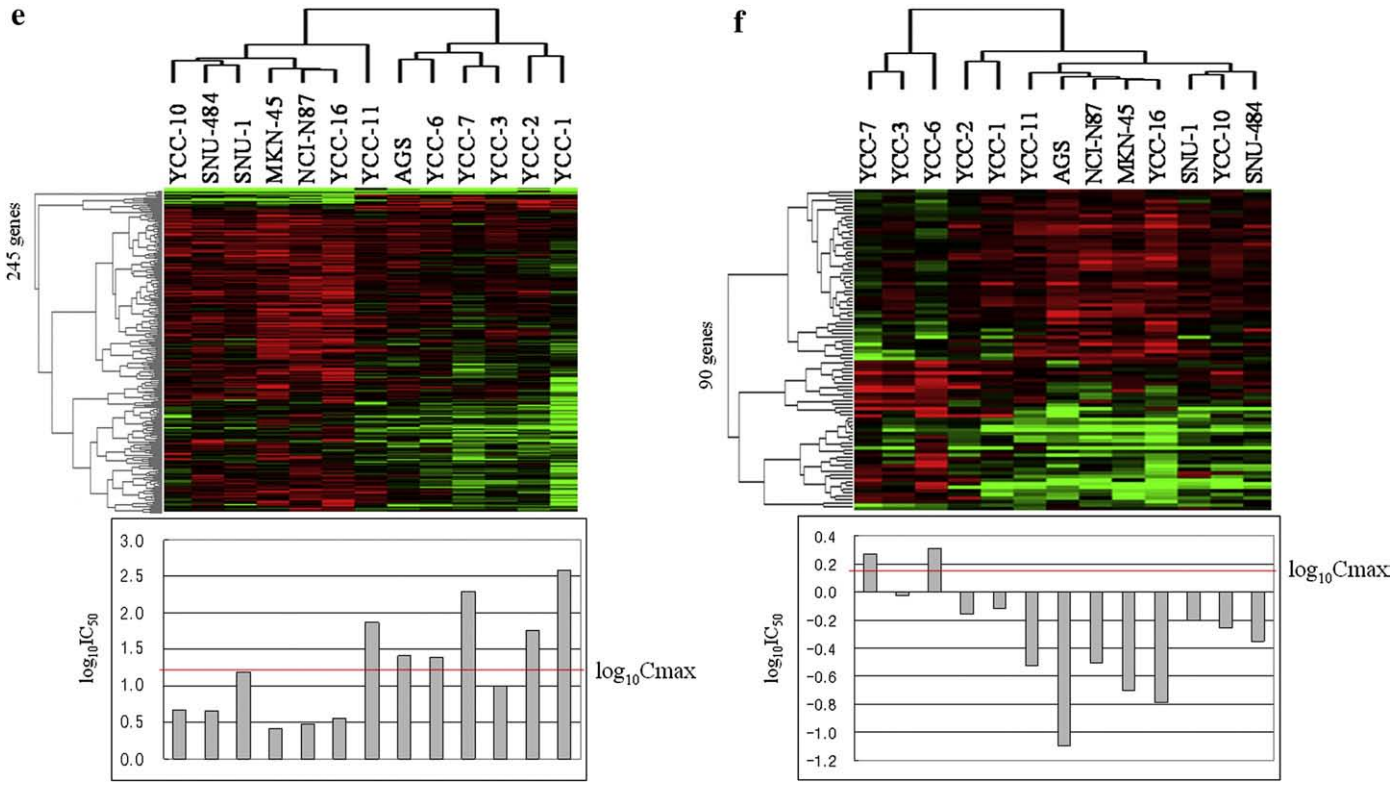

Fig. 3 (*continued*).

presumed to be common chemosensitivity-related genes in gastric cancer. Hierarchical clustering with these 85 genes and the anticancer drugs showed negative correlation (chemo-sensitive) and positive correlation (chemo-resistant) (Fig. 2a). Then, the expression patterns of these genes were examined in the same order as in Fig 2a for the 13 gastric cancer cell lines, with green representing down-regulated expression, and red representing up-regulated expression (Fig. 2b). The expression patterns of the 85 genes largely classified the cell lines into three groups, representing cell lines that are resistant, sensitive, and of intermediate sensitivity, to multiple drugs. It represents that genes with positive correlation were over-expressed in common chemo-resistant cell lines, and negatively correlated genes were over-expressed in common chemo-sensitive cell lines. We also performed the multidimensional scaling (MDS) (Fig. 2c) with the expression of the 85 genes, and confirmed that the common chemo-resistant and -sensitive cell lines were well separated.

Selection of genes associated with drug-specific chemosensitivity

Genes with a correlation coefficient in the GC matrix of ≥|0.6| were subjected to two-class SAM analysis with the gastric cancer cell lines to identify genes that could classify the cell lines as chemo-resistant or-sensitive for each drug. The hierarchical clustering represented well-dichotomized genes that were strongly associated with chemo-resistance or-sensitivity to specific drugs. First, we could not analyze the cyclophosphamide, leucovorin, methotrexate, because the IC_{50}s of cell lines in these drugs were higher than their Cmax, as mentioned above. Using this algorithm, we could select genes associated with chemosensitivity for cisplatin, irinotecan, paclitaxel, docetaxel, 5-FU and doxorubicin, commonly used in gastric cancer, with low false discovery rate (FDR, 0–0.2) (Fig. 3a–f). Each FDR was determined by SAM in the result of gene selection. Table 2 shows the top 20 ranked genes associated with chemosensitivity for each drug. The results for the remaining 7 drugs are summarized in Supplementary figure. All of the selected genes associated with drug-specific chemosensitivity are listed in a Supplementary data file 4.

Microarray validation with quantitative RT-PCR

To validate the microarray results, RNA expression of randomly selected 2 genes, SOX9 and PTGER2, was performed using qRT-PCR analysis in the 13 gastric cancer cell lines. Microarray and qRT-PCR data were well correlated to each other with the Pearson correlation coefficient of 0.92 and 0.69 in SOX9 and PTGER2, respectively (Fig. 4a). Linear regression model also showed that these two types of measurements had significant relation, $p<0.05$ (Fig. 4a). In the microarray data SOX9 and PTGER2 were selected as down-regulated genes in resistant cell lines to gemcitabine and 5-FU, respectively, and that were correctly confirmed in qRT-PCR results (Fig. 4b).

Discussion

In this study we analyzed the gene expression profiles of 13 human gastric cancer cell lines, and compared them to chemosensitivity data from the same cell lines. Two similar studies have previously been carried out using *in vitro* and *in vivo* animal model system [2,4]. Scherf et al. constructed the first pharmacogenomic database using 60 cell lines, 118 drugs and 1376 genes [2]. However, our study has different aims, namely to: 1) expand the number of genes to 17,000 for more genome-wide evaluation, 2) focus on gastric cancer cell lines, and 3) evaluate recently developed anticancer drugs which are commonly used for gastric cancer in the clinic.

Few studies have investigated genes associated with multi-drug chemosensitivity using anticancer agents currently used in gastric cancer. Accordingly, we would like to select gene sets related to chemosensitivity for a large number of anticancer agents. Clinically, such genes may be used to identify patients who are refractory to currently used drugs, and thus need to enter clinical trials with newly

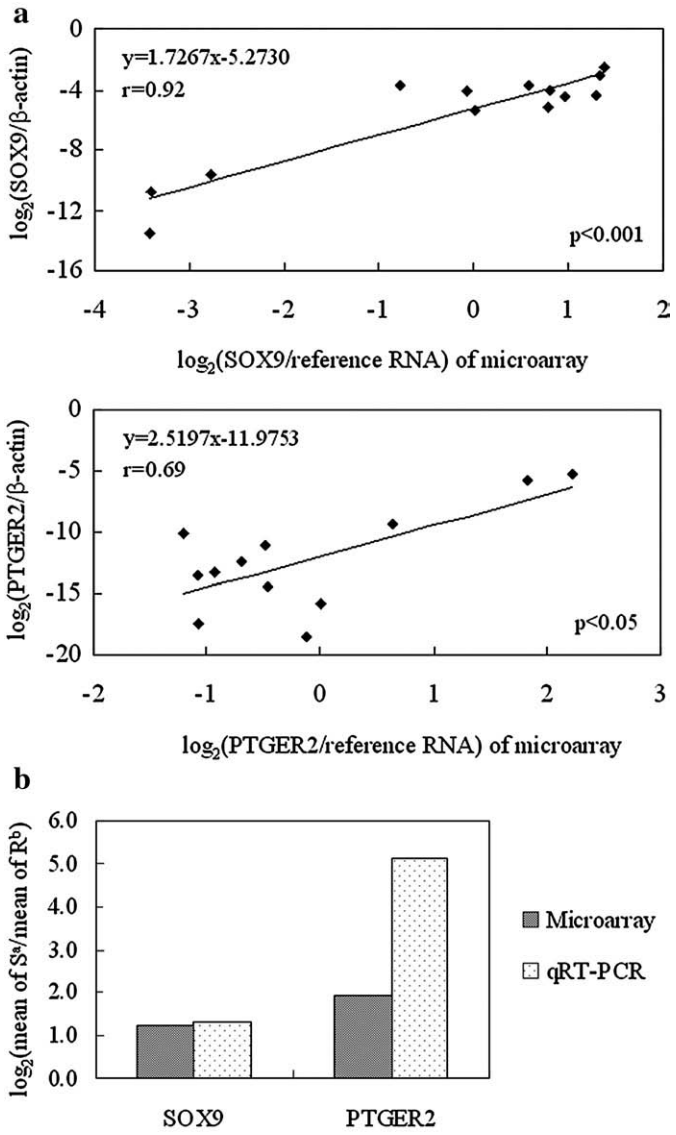

Fig. 4. Microarray validation with quantitative RT-PCR. (a) Linear regression model for microarray and qRT-PCR $\log_2$ ratio of SOX9 (upper) and PTGER2 (lower) genes in 13 gastric cancer cell lines. The regression formulas, R^2 values and Pearson correlation coefficient (r) are shown on the left upper corner, and p-values for linear regression are on the right lower corner of each panel. In the qRT-PCR result of PTGER2, gene expression in SNU-1 cell line was not detected, so the linear regression was done with 12 cell lines. (b) Relative fold changes of SOX9 and PTGER2 between resistant and sensitive groups to gemcitabine and 5-FU, respectively. The patterns of microarray and qRT-PCR were well matched. [a], sensitive cell lines. [b], resistant cell lines.

developed agents. Alternatively, we may select patients who could benefit from the narrow therapeutic window of current chemotherapy, and for whom drugs with low toxicity can be the best choice.

For the selection of genes associated with common chemoresistance or -sensitivity, we selected genes with a correlation coefficient greater than |0.40| for 11 of 16 (≥70%) anticancer drugs, and identified 85 genes (37 positively correlated and 48 negatively correlated, Table 1). A correlation coefficient of |0.40| was the value at which we could find reasonable number of chemosensitivity-related genes. When we applied the grouping method of Staunton et al. [5], cell lines with common resistance were positioned at the upper outlier area of IC_{50}±0.6 standard deviation, while sensitive cell lines were positioned at the lower outlier (Supplementary Table 1). Because we selected only genes with a relatively high correlation, some genes previously identified as being associated with chemosensitivity but which had a relatively low correlation, were not selected. We did, however, confirm previously reported genes such as MAD2L1, TXN and FAS, which showed a high correlation with the drugs studied. Among them, MAD2 is known to be involved in cellular mitotic arrest, and is associated with multi-drug resistance in gastric cancer [17–19]. We also found that MAD2L1 correlated negatively with most of anti-tubule agents including etoposide, vinblastine, carboplatin, methotrexate, paclitaxel, and vincristine (r=−0.75, −0.52, −0.51, −0.42, −0.39, and −0.37, respectively). In addition, as Pranav Sinha et al., reported that thioredoxin (TXN) was over-expressed in a multi-drug resistant gastric cancer cell line [20], we found that TXN had positive correlation with topotecan, irinotecan and cisplatin (r=0.65, 0.51 and 0.40, respectively). Over-expression of the FAS gene sensitized drug resistant gastric cancer cells to chemotherapeutic drugs [21,22], and is known to be associated with inrinotecan sensitivity [23,24]. In our data, FAS had significant negative correlations with irinotecan, topotecan, mitomycin, and gemcitabine (r=−0.88, −0.80, −0.57 and −0.41, respectively, Fig. 5). These results support the reliability of our GC-matrix based pharmacogenomic database with much wider

Table 1
The 85 genes related to common chemosensitivity

Positively correlated genes		Negatively correlated genes	
GenBank no.	Symbol	GenBank no.	Symbol
NM_015367	BCL2L13	NM_001643	APOA2
NM_024952	C14orf159	NM_020375	C12orf5
NM_017759	FLJ20309	NM_030568	KHDC1
NM_144600	C16orf63	NM_022102	CCDC90A
NM_022158	FN3K	NM_019084	CCNJ
NM_005814	GPA33	NM_001799	CDK7
NM_000841	GRM4	NM_021195	CLDN6
NM_033445	HIST3H2A	NM_018403	DCP1A
NM_032303	HSDL2	NM_022365	DNAJC1
NM_024013	IFNA1	NM_001387	DPYSL3
NM_016027	LACTB2	NM_005088	SFRS17A
NM_002969	MAPK12	NM_022336	EDAR
NM_005918	MDH2	NM_013241	FHOD1
NM_024060	AHNAK	NM_025139	ARMC9
NM_022474	MPP5	NM_014366	GNL3
NM_002514	NOV	NM_004832	GSTO1
NM_005047	PSMD5	NM_001524	HCRT
NM_002842	PTPRH	NM_016185	HN1
NM_002888	RARRES1	NM_014804	KIAA0753
NM_014059	C13orf15	NM_012311	KIN
NM_017585	SLC2A6	NM_058169	LOH12CR1
NM_003098	SNTA1	NM_144565	DUOXA1
NM_020777	SORCS2	NM_017830	OCIAD1
NM_054114	TAGAP	NM_022375	OCLM
NM_013259	TAGLN3	NM_020190	OLFML3
NM_003234	TFRC	NM_148961	OTOS
NM_020131	UBQLN4	NM_006206	PDGFRA
NM_031407	HUWE1	NM_005777	RBM6
NM_032050	PATZ1	NM_013349	NENF
XM_096883	–	NM_012244	SLC7A8
XM_040195	–	NM_005985	SNAI1
XM_170868	–	NM_003201	TFAM
XM_098154	–	NM_033285	TP53INP1
XM_173309	–	NM_032236	USP48
NM_020224	–	NM_007268	VSIG4
XM_032852	–	NM_144621	ZBTB8
XM_047287	–	XM_060462	–
		XM_113998	–
		NM_032739	–
		NM_032715	–
		NM_005758	–
		XM_173080	–
		XM_113861	–
		XM_092919	–
		XM_047196	–
		XM_066500	–
		XM_171129	–
		XM_059017	–

–:Unknown gene.

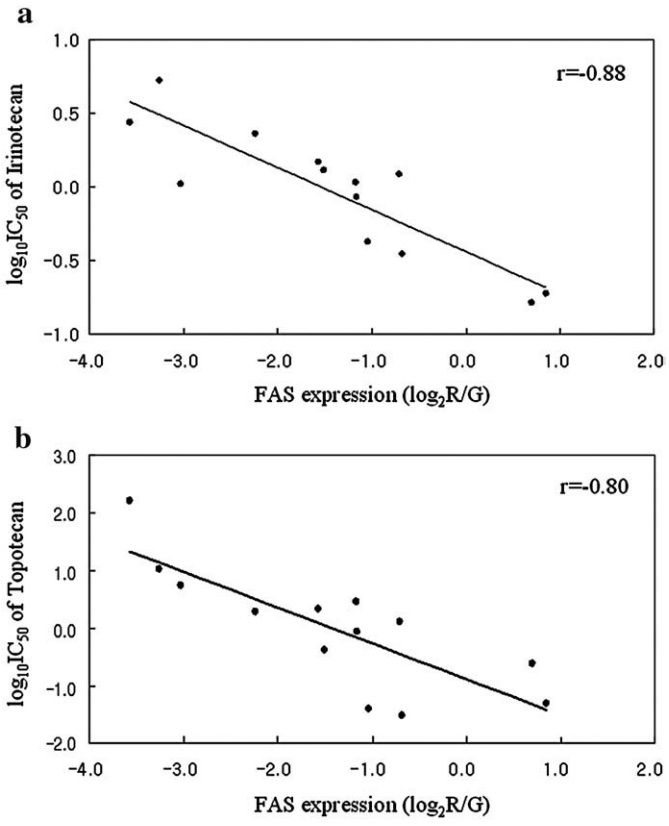

Fig. 5. Relationship between FAS expression and chemosensitivity to irinotecan and topotecan in gastric cancer cell lines. (a) Relationship between FAS and irinotecan (−0.88 correlation coefficient). (b) Relationship between FAS and topotecan (−0.80 correlation coefficient).

spectrum of chemo-sensitive or-resistant genes in more diverse chemotherapeutic agents.

While selecting the chemosensitivity genes related to each specific drug, we faced a problem that we couldn't select the genes with simple SAM analysis only using G-matrix. It can be explained as there are too many genes without influence on the chemosensitivity, those noisy genes might disturb the SAM in selecting significant genes. So, we introduced the GC-matrix into the process of gene selection, and could filter out the genes which had relatively low correlation coefficient than other genes. Finally, we could select the chemosensitivity related genes with two step analysis method, GC-matrix and SAM, differently from previous studies. We confirmed previously reported chemosensitivity genes encoding targets of anticancer drugs, which had a moderate correlation coefficient as in the previous studies, such as TP with 5-FU (r=0.28), UPP1 with 5-FU (r=−0.34), and TOP2A,-B with etoposide (r=−0.35, −0.37, respectively). In addition, some genes with unknown function could be found to have similar correlation coefficients to those reported previously [4], e.g., FAM38A with 5-FU (r=0.42) and HMGCL, HSD17B2 with cisplatin (r=−0.51, 0.49, respectively). In our study, we applied more strict criteria for gene selection with higher correlation coefficient more than |0.6|, resulting in many novel genes. In the case of irinotecan, some genes associated with apoptosis and cell cycle arrest were over-expressed in sensitive cell lines, including FAS (r=−0.88), AIF-1 (r=−0.73), FKBP1A (r=−0.66), TNFRSF1B (r=−0.64), EDAR (r=−0.61), and BBC3 (r=−0.67). Interestingly, FAS (apoptosis), and AIF1 (cell cycle arrest) were also strongly correlated with topotecan, a drug with a similar mechanism of action as iritotecan, (r=−0.80, −0.75, respectively).

Increased expression of thymidylate synthetase (TYMS), which inactivates 5-FU to 5-fluoro-dUMP, was reported to correlate with gastric cancer cell resistance to 5-FU [4,25,26]. In our study, however, TYMS was not selected in the 5-FU related gene set and the correlation coefficient between TYMS and 5-FU was as low as 0.038. In tracking of this result, we found that all of the expression of TYMS in 13 cell lines was similarly low regardless of 5-FU chemosensitivity. Possible reasons might be considered as below; 1) The overexpression of TYMS may be occurred after the drug treatment, 2) The function of TYMS may be activated with other 5-FU related genes, 3) Relationship between TYMS and other genes associated with leucovorin may influence the expression of TYMS, and 4) When a mutation of TYMS is occurred, the expression of TYMS may be over-expressed. We then, therefore, focused on the 245 selected genes that were associated with sensitivity to 5-FU (Fig. 3e). Of these, only 15 genes had a positive correlation (Table 2), and which were relatively over-expressed in 5-FU resistant cell lines, suggesting the association with 5-FU resistance.

To evaluate another applicability of our pharmacogenomic database, we compared the GC-matrix of similar anti-tubule agents, paclitaxel and docetaxel. We identified independent sets of genes associated with paclitaxel or docetaxel chemosensitivity, i.e., there were no genes that were selected for both paclitaxel and docetaxel, although the functions of selected genes were similar for the two drugs (data not shown). For example, many genes related to cell adhesion, cell proliferation and proteolysis were selected in both drugs, although the specific genes were different. It suggested that the drugs may have different mechanisms of action and providing a possible explanation for the difference in efficacy and toxicity spectrum of these drugs. In this way, we can compare specific genes associated with drugs that have similar mechanisms of action, or drugs that are used for combination chemotherapy.

This study focused on the intrinsic susceptibility of gastric cancer to various anticancer drugs with respect to potential application as predictive markers before introducing chemotherapy. Therefore, in contrast with several researches which used cancer cell lines with acquired drug-resistance [27] or immediately following drug treatment [28], we used untreated gastric cancer cell lines. This is a first report that revealed the relation between genes and anticancer drugs currently used for gastric cancer via high-throughput method.

We identified several sets of chemosensitivity-related genes for a specific drug or multiple drugs based on the GC-matrix of a gastric cancer cell line database. These genes may act as predictive markers for chemosensitivity in chemo-naive gastric cancer patients following functional analysis and clinical validation.

Materials and methods

Cell lines and culture

Two human gastric cancer cell lines, AGS and NCI-N87, were obtained from the American Type Culture Collection (Rockville, Maryland, USA). MKN-45 and the SNU-series (SNU-1, SNU-484) were obtained from the Japanese Cancer Research Resources Bank and Korean Cell Line Bank, respectively. Eight cell lines, YCC-1, YCC-2, YCC-3, YCC-6, YCC-7, YCC-10, YCC-11 and YCC-16, were established from Korean gastric cancer patients at the Cancer Metastasis Research Center (CMRC, Yonsei University College of Medicine, Seoul, Korea). Cells were cultured under conditions provided by the manufacturer, and were incubated at 37 °C in a 5% CO_2 humidified atmosphere and the media replaced every 3 days.

RNA preparation

Total RNA was extracted from each cell line using TRIzol reagent (Invitrogen, Carlsbad, CA, USA) according to the manufacturer's instructions. The Yonsei reference RNA (Cancer Metastasis Research

Table 2
The top ranked genes positively and negatively associated with drug specific chemosensitivity for 6 drugs

Cisplatin		Irinotecan		5-FU		Doxorubicin		Paclitaxel		Docetaxel	
Symbol	*r*	Symbol	*r*	Symbol	*r*	Symbol	*r*	Symbol	*r*	Symbol	*r*
INSL4	0.8415	–	0.7474	AURKA	0.7632	TAF7L	0.8077	–	0.8355	–	0.8485
AKAP8L	0.8409	PHKA1	0.7398	HIST1H2AM	0.7623	–	0.8063	INO80	0.8087	MGST2	0.8275
ZNF613	0.8243	–	0.7282	ASPM	0.7333	INHBB	0.7547	INPP4B	0.8002	–	0.8274
PPHLN1	0.8056	PHGDH	0.6883	DDX58	0.7105	TBX3	0.7406	STEAP4	0.7554	C14orf108	0.8102
DSC2	0.7924	PFN2	0.6799	PI3	0.7086	CD151	0.7323	BRE	0.7501	–	0.8074
NKAP	0.7913	ARHGEF10	0.657	TRIP10	0.6841	DNAJC11	0.7288	MTHFD1L	0.7501	SLC6A13	0.807
NIT1	0.7847	–	0.6444	DHRS2	0.6661	GBP1	0.7266	SH3TC2	0.7451	IGSF8	0.8042
HOXB5	0.7817	CTSL1	0.6419	IFI6	0.6621	BIRC2	0.7185	MRPL33	0.6923	DHRS1	0.801
C1orf177	0.7761	PYGL	0.6051	AXL	0.6602	TSPAN2	0.6996	PLD1	0.6786	–	0.7926
KLF12	0.7705	TNFRSF1B	−0.6433	SLC39A1	0.6549	GPR175	0.6902	–	0.6707	–	0.7662
ALPL	0.7626	ARFGAP1	−0.644	ATP1A1	0.6469	BIRC3	0.6801	IFNGR1	0.6563	–	0.7656
HMX1	0.7624	DTNBP1	−0.6599	BAIAP2	0.6261	–	0.6729	TNFSF13B	0.6551	GFOD1	0.76
–	0.7566	ZNF232	−0.6612	TRIM29	0.6209	IRF1	0.6671	PDE9A	0.6376	C6orf192	0.7499
SARS	0.7557	FKBP1A	−0.6647	THBS1	0.6208	PCDHB3	0.6635	CD302	0.63	FAM129B	0.7359
MERTK	0.7555	BBC3	−0.6684	OLR1	0.6206	HSPA1B	0.6631	NOV	0.6293	ABHD4	0.7338
NEU1	0.7535	C19orf53	−0.6693	–	−0.7451	SH3GL3	0.6627	–	0.625	C14orf94	0.729
–	0.7521	–	−0.673	LZTFL1	−0.7485	CCND3	0.6596	TTLL7	0.6246	RABGGTA	0.7194
MARS	0.7509	–	−0.6922	DEFB119	−0.7495	ZC3HDC1	0.6579	PCDH7	0.6185	–	0.7149
CDC26	0.7489	LOC197336	−0.6936	SCYE1	−0.7504	UBA7	0.6564	LGALS3BP	0.6125	MYLIP	0.7101
–	0.7469	FLI1	−0.7005	TSG101	−0.7518	CBX6	0.6458	HERC6	0.6048	STXBP6	0.702
ACCN3	−0.7637	–	−0.7011	–	−0.7545	TFAP2B	−0.6975	AFF3	−0.725	SDC3	−0.6777
MYBL1	−0.7665	ATP13A1	−0.7173	FLJ22318	−0.7575	–	−0.6988	SERPINB6	−0.7289	MAP1LC3B	−0.6778
–	−0.7717	–	−0.7194	LONP1	−0.7588	CASC1	−0.7064	CBX1	−0.7362	RAB28	−0.6851
KCNMA1	−0.7758	AIF1	−0.7281	–	−0.7608	–	−0.7113	–	−0.7368	ZBTB12	−0.7007
–	−0.7786	NMUR2	−0.7296	SENP6	−0.7636	–	−0.7122	LIMS2	−0.7384	APOL3	−0.7082
BMP1	−0.7858	C1GALT1	−0.7319	ZNF277	−0.7674	ZNF589	−0.7168	GPC4	−0.7421	RBAK	−0.7172
ASAH2	−0.7862	RFX1	−0.7338	PRPSAP1	−0.7693	–	−0.7185	KIFAP3	−0.7543	MRPS6	−0.7186
CXCR4	−0.7864	HOXC5	−0.7481	FAM167B	−0.7771	–	−0.719	LTK	−0.7612	CD68	−0.7258
OSR2	−0.7883	–	−0.749	BPTF	−0.7798	ZBTB25	−0.7234	DLEU2	−0.7616	–	−0.741
PARP3	−0.7947	–	−0.7587	HESX1	−0.7915	–	−0.7387	–	−0.7649	TRAPPC3	−0.7559
RBM30	−0.7972	–	−0.7718	PTER	−0.7928	BAZ2B	−0.7396	LRP4	−0.7667	SLC5A3	−0.7661
KCNAB2	−0.8016	UGT2B4	−0.7761	AKR1A1	−0.7982	IGJ	−0.7641	SMAD5	−0.7677	–	−0.7698
–	−0.8077	NFE2	−0.7819	ALOX15B	−0.8079	ENOSF1	−0.7649	PNRC2	−0.7709	EPN2	−0.7793
Rgr	−0.8112	IGHMBP2	−0.7915	–	−0.8205	METTL4	−0.789	CCR3	−0.7764	C6orf129	−0.8042
PTTG1IP	−0.8132	ROR1	−0.8005	SLC22A17	−0.8238	–	−0.8144	PBX3	−0.7836	IQCG	−0.8055
–	−0.8211	PTS	−0.82	POLL	−0.8259	–	−0.8277	–	−0.7955	ZMYM6	−0.8146
–	−0.8328	LDB1	−0.8301	RICS	−0.8391	FKSG2	−0.8414	ADAMTS19	−0.7967	GBF1	−0.8296
–	−0.8431	HYI	−0.8304	BAAT	−0.8581	PCDHGC3	−0.856	ABCG4	−0.8075	MED8	−0.8344
NRP2	−0.8713	FAS	−0.8762	IRF2	−0.8639	–	−0.8645	NDUFV3	−0.8178	MAP4K2	−0.863
PLCB2	−0.8824	C10orf11	−0.8832	KIAA1737	−0.8699	PRND	−0.9166	ZNF564	−0.8301	–	−0.877

–: Unknown gene.
r: Pearson correlation coefficient.

Center, Seoul, Korea) was prepared by pooling equivalent amounts of total RNA from the following 11 human cancer cell lines: YCC-B1 (breast cancer), HCT-116 (colon cancer), SK-HEP-1 (liver cancer), A549 (lung cancer), HL-60 (acute promyelocyte leukemia), MOLT-4 (acute lymphoblastic leukemia), HeLa (cervical cancer), Caki-2 (kidney cancer), T98G (glioblastoma), HT1080 (fibrosarcoma) and YCC-3 (gastric cancer) [29]. The quantity and quality of RNA were confirmed by a ND-1000 spectrophotometer (NanoDrop Technologies, USA) and gel electrophoresis.

Oligonucleotide microarray

Oligonucleotide microarray analysis was performed using a human oligo chip (CMRC-GT, Seoul, Korea) containing 22,740 oligonucleotide probes of 70 bases with a reference design. The test samples (RNA from each gastric cancer cell line) were labeled with Cy5 and individually co-hybridized with the Cy3-labeled reference RNA (CMRC, Seoul, Korea). For further analysis, raw Cy5/Cy3 data were $\log_2$-transformed. Systemic errors were corrected by normalization using intensity dependent, within-print, tip normalization based on the Lowess function. After normalization, genes with more than one missing value in all experiments were filtered, no missing proportion (NMP) 100%. The values of repeated genes were adjusted by S-Plus 2000 software (Insightful, Seattle, WA, USA). For the reliability of experiments, YCC-10 was performed triplicate.

Chemosensitivity assay

Growth inhibition was measured in 13 human gastric cancer cell lines with 16 anticancer drugs (5-fluorouracil, irinotecan, topotecan, doxorubicin, etoposide, mitomycin, docetaxel, paclitaxel, cisplatin, carboplatin, gemcitabine, vinblastine, vincristine, leucovorin, cyclophosphamide and methotrexate,) provided by Yonsei Cancer Center (Seoul, Korea) using the MTT (3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide, Sigma, Saint Louis, MO, USA) assay. The drug concentration at which 50% of cancer cells survived (IC_{50}, μM) was calculated using Calcusyn software (Biosoft, Cambridge, UK). Since there was wide variation in the scale of data points for different drugs, the IC_{50} was transformed into a $\log_{10}$ scale. Among the 16 anticancer drugs, leucovorin and cyclophosphamide were used as a negative control because leucovorin itself does not have any anti-tumor activity and cyclophosphamide need *in vivo* activation for the anti-tumor activity.

GC-matrix analysis

The degree of similarity between the G-matrix and C-matrix was calculated using the Pearson correlation coefficient, *r*, calculated by MATLAB software (The MathWorks, Inc) as follows:

$$r = \frac{\sum (x_i - x_m)(y_i - y_m)}{\sqrt{\sum (x_i - x_m)^2 (y_i - y_m)^2}}$$

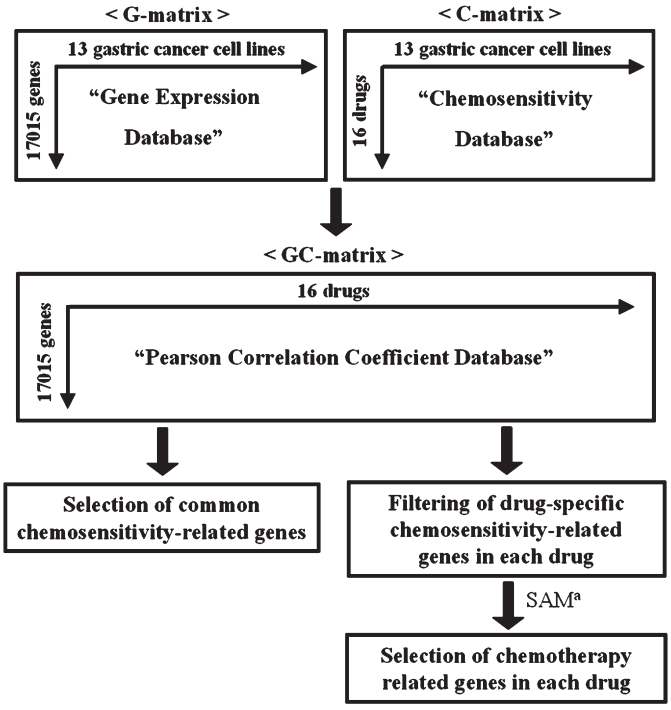

Fig. 6. A schematic of the study design. [a], significance analysis of microarrays.

Where x_i denotes the $\log_2$ gene expression ratio ($\log_2$Cy5/Cy3) of gene × in cell *i*, and y_i denotes the $\log_{10}$ chemosensitivity ($\log_{10}IC_{50}$) of cell *i* to drug *y*. x_m represents the mean of the $\log_2$ gene expression ratio of gene *x*, and y_m represents the mean chemosensitivity of the drug.

Data analysis

A schematic of the study design is shown in Fig. 6. The gastric cancer cell lines were divided into chemo-sensitive and chemo-resistant groups based on the IC_{50} following the cut-off value of the clinically achievable maximum concentration (Cmax, Supplementary Table 2) of each drug. First, to identify the common chemosensitivity-related genes, we selected the genes with a correlation coefficient greater than |0.40| in 11 of 16 anticancer drugs using GC-matrix. Next, to identify the chemosensitivity-related genes for each specific drug, the GC-matrix was used to select the genes with a correlation coefficient greater than |0.60| for each drug. We finally determined the significant genes which could divide the cell lines into chemo-sensitive and -resistant groups using two-class significance analysis of microarrays (SAM) [30] with selected genes.

Hierarchical clustering analysis was performed with Cluster (Eisen Lab, http://rana.lbl.gov/EisenSoftware.htm) and the resulting dendrogram was visualized using TreeView software (Eisen Lab). Clustering was done by complete linkage algorithm with uncentered correlation. The distance of each cluster represents correlation between two clusters. Multidimensional scaling (MDS) was done by using the BRB-Arraytools version 3.3.0 software package (http://linus.nci.nih.gov/BRB-ArrayTools.html) [31].

Annotation of the selected genes was performed using the Database for Annotation, Visualization and Integrated Discovery (DAVID) (http://apps1.niaid.nih.gov/david) and the Stanford Online Universal Resource for Clones and Expressed sequence tags (SOURCE) (http://source.stanford.edu/cgi-bin/source/sourceSearch).

Quantitative RT-PCR

SOX9 and PTGER2 were randomly selected for validation of the microarray data. Quantitative RT-PCR (qRT-PCR) was performed on 13 gastric cancer cell lines. In brief, 4μg of total RNA from each sample was reverse-transcribed using SuperScript II Reverse Transcriptase (Invitrogen, Carlsbad, CA, USA). Two hundred nanograms of synthesized cDNA were PCR amplified using QuantiTect SYBR Green PCR (QIAGEN, Valencia, CA, USA). Each reaction was run in duplicate on a Stratagene MX3005P (Stratagene, La Jolla, CA, USA). Expression values for each gene were determined using a standard curve constructed from Human Genomic DNA (Promega, Madison, WI, USA). The house-keeping gene ACTB was selected for normalization and the standard curve. Non-template-control wells without cDNA were included as negative controls. The primer sets for PCR amplification were designed as follows: SOX9-Sense: 5′-TTTCCAAGACACAAACATGA-3′, SOX9-Antisense: 5′-AAAGTCCAGTTTCTCGTTGA-3′, PTGER2-Sense: 5′-GCTATCATGACCATCACCTT-3′, PTGER2-Antisense: 5′-TGTGTTGCATCTTGTGTTCT-3′. To compare gene expressions between microarray and qRT-PCR, we calculated Pearson correlation coefficient, and built a linear regression model based on $\log_2$ ratio of real-time PCR.

Acknowledgments

Sun Young Rha's research was supported by the Korea Science and Engineering Foundation (KOSEF) through the Cancer Metastasis Research Center (CMRC) at Yonsei University College of Medicine. Hyun Cheol Chung's research was supported by a grant of the Korea Health 21 R&D Project, Ministry of Health & Welfare, the Republic of Korea (0405-BC01-0604-0002).

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.ygeno.2008.08.002.

References

[1] D.T. Ross, et al., Systematic variation in gene expression patterns in human cancer cell lines, Nat. Genet. 24 (2000) 227–235.
[2] U. Scherf, et al., A gene expression database for the molecular pharmacology of cancer, Nat. Genet. 24 (2000) 236–244.
[3] S. Dan, et al., An integrated database of chemosensitivity to 55 anticancer drugs and gene expression profiles of 39 human cancer cell lines, Cancer Res. 62 (2002) 1139–1147.
[4] H. Zembutsu, et al., Genome-wide cDNA microarray screening to correlate gene expression profiles with sensitivity of 85 human cancer xenografts to anticancer drugs, Cancer Res. 62 (2002) 518–527.
[5] J.E. Staunton, et al., Chemosensitivity prediction by transcriptional profiling, Proc. Natl. Acad. Sci. U. S. A. 98 (2001) 10787–10792.
[6] J.N. Weinstein, et al., An information-intensive approach to the molecular pharmacology of cancer, Science 275 (1997) 343–349.
[7] S. Pyrhonen, T. Kuitunen, P. Nyandoto, M. Kouri, Randomised comparison of fluorouracil, epidoxorubicin and methotrexate (FEMTX) plus supportive care with supportive care alone in patients with non-resectable gastric cancer, Br. J. Cancer 71 (1995) 587–591.
[8] A.M. Murad, et al., Modified therapy with 5-fluorouracil, doxorubicin, and methotrexate in advanced gastric cancer, Cancer 72 (1993) 37–41.
[9] B. Glimelius, K. Hoffman, U. Haglund, O. Nyren, P.O. Sjoden, Initial or delayed chemotherapy with best supportive care in advanced gastric cancer, Ann. Oncol. 5 (1994) 189–190.
[10] H.Y. Wang, et al., Assessing unmodified 70-mer oligonucleotide probe performance on glass-slide microarrays, Genome Biol. 4 (2003) R5.
[11] E. Southern, K. Mir, M. Shchepinov, Molecular interactions on microarrays, Nat. Genet. 21 (1999) 5–9.
[12] A. Watson, A. Mazumder, M. Stewart, S. Balasubramanian, Technology for microarray analysis of gene expression, Curr. Opin. Biotechnol. 9 (1998) 609–614.
[13] X. Zhao, S. Nampalli, A.J. Serino, S. Kumar, Immobilization of oligodeoxyribonucleotides with multiple anchors to microchips, Nucleic Acids Res. 29 (2001) 955–959.
[14] D.R. Call, D.P. Chandler, F. Brockman, Fabrication of DNA microarrays using unmodified oligonucleotide probes, Biotechniques 30 (2001) 368–372 374, 376 passim.
[15] M.D. Kane, et al., Assessment of the sensitivity and specificity of oligonucleotide (50mer) microarrays, Nucleic Acids Res. 28 (2000) 4552–4557.
[16] A. Barczak, et al., Spotted long oligonucleotide arrays for human gene expression analysis, Genome Res. 13 (2003) 1775–1785.
[17] C.W. Wu, C.W. Chi, T.S. Huang, Elevated level of spindle checkprotein MAD2 correlates with cellular mitotic arrest, but not with aneuploidy and clinicopathological characteristics in gastric cancer, World J. Gastroenterol. 10 (2004) 3240–3244.
[18] F. Yin, W.H. Hu, T.D. Qiao, D.M. Fan, [Multidrug resistant effect of alternative splicing form of MAD2 gene-MAD2beta on human gastric cancer cell], Zhonghua Zhong Liu Za Zhi 26 (2004) 201–204.

[19] K. Tanaka, et al., Mitotic checkpoint protein hsMAD2 as a marker predicting liver metastasis of human gastric cancers, Jpn. J. Cancer Res. 92 (2001) 952–958.
[20] P. Sinha, et al., Increased expression of annexin I and thioredoxin detected by two-dimensional gel electrophoresis of drug resistant human stomach cancer cells, J. Biochem. Biophys. Methods 37 (1998) 105–116.
[21] B. Xiao, et al., Transduction of Fas gene or Bcl-2 antisense RNA sensitizes cultured drug resistant gastric cancer cells to chemotherapeutic drugs, World J. Gastroenterol. 4 (1998) 421–425.
[22] F. Yin, et al., Suppression of P-gp induced multiple drug resistance in a drug resistant gastric cancer cell line by overexpression of Fas, World J. Gastroenterol. 6 (2000) 664–670.
[23] L. Catley, et al., Proteasomal degradation of topoisomerase I is preceded by c-Jun NH2-terminal kinase activation, Fas up-regulation, and poly(ADP-ribose) polymerase cleavage in SN38-mediated cytotoxicity against multiple myeloma, Cancer Res. 64 (2004) 8746–8753.
[24] U. McDermott, et al., Effect of p53 status and STAT1 on chemotherapy-induced, Fas-mediated apoptosis in colorectal cancer, Cancer Res. 65 (2005) 8951–8960.
[25] P.G. Johnston, et al., Thymidylate synthase gene and protein expression correlate and are associated with response to 5-fluorouracil in human colorectal and gastric tumors, Cancer Res. 55 (1995) 1407–1412.
[26] D. Salonga, et al., Colorectal tumors responding to 5-fluorouracil have low gene expression levels of dihydropyrimidine dehydrogenase, thymidylate synthase, and thymidine phosphorylase, Clin. Cancer Res. 6 (2000) 1322–1327.
[27] H.C. Kang, et al., Identification of genes with differential expression in acquired drug-resistant gastric cancer cells using high-density oligonucleotide microarrays, Clin. Cancer Res. 10 (2004) 272–284.
[28] J.S. Park, et al., Identification of novel genes associated with the response to 5-FU treatment in gastric cancer cell lines using a cDNA microarray, Cancer Lett. 214 (2004) 19–33.
[29] T.M. Kim, et al., Determination of genes related to gastrointestinal tract origin cancer cells using a cDNA microarray, Clin. Cancer Res. 11 (2005) 79–86.
[30] V.G. Tusher, R. Tibshirani, G. Chu, Significance analysis of microarrays applied to the ionizing radiation response, Proc. Natl. Acad. Sci. U. S. A. 98 (2001) 5116–5121.
[31] R. Rouzier, et al., Breast cancer molecular subtypes respond differently to preoperative chemotherapy, Clin. Cancer Res. 11 (2005) 5678–5685.